# The accuracy is:

- a. The number of correct significant digits in approximating some quantity.
- b. The number of digits with which a number is expressed.
- o. None of the above.

Given two random variables X and Y, Bayes Theorem implies that  $p(y|x)=rac{p(x|y)p(y)}{p(x)}$  where:

- $\bigcirc$  a. p(x|y) is called prior distribution on x.
- $\bigcirc$  b. p(x|y) is called posterior distribution on y.
- $\bigcirc$  c. p(x|y) is called likelihood on y.

Given two random variables X and Y, Bayes Theorem implies that  $p(y|x)=rac{p(x|y)p(y)}{p(x)}$  where:

- $\bigcirc$  a. p(y|x) is called prior distribution on x.
- $\bigcirc$  b. p(y|x) is called likelihood on y.
- $\bigcirc$  c. p(y|x) is called posterior distribution on y.

If  $f:\mathbb{R}^2 o\mathbb{R}$ ,  $f(x_1,x_2)=x_1e^{x_2}$ ,  $g:\mathbb{R} o\mathbb{R}^2$ ,  $g(t)=(e^t,t)$ , then, if h(t)=f(g(t)):

- $\bigcirc$  a.  $h'(t)=te^t$ .
- $\bigcirc$  b.  $h'(t)=2e^{2t}$ .
- 0 c.  $h'(t) = e^{2t}(t+1)$ .

If  $f:\mathbb{R}^2 o\mathbb{R}$ ,  $f(x_1,x_2)=x_1^2+x_1x_2$ ,  $g:\mathbb{R} o\mathbb{R}^2$ ,  $g(t)=(\sin(t),\cos(t))$ , then, if h(t)=f(g(t)):

- $\bigcirc$  a.  $h'(t) = \sin(2t) \sin^2(t)$ .
- $\circ$  b.  $h'(t) = \sin(t) \sin^2(2t)$ .
- $\odot$  c.  $h'(t) = \sin(t)\cos(t) \sin^2(t)$ .

If  $f:\mathbb{R}^2 o \mathbb{R}$ ,  $f(x_1,x_2)=x_1^2+x_2^2$ ,  $g:\mathbb{R}^2 o \mathbb{R}^2$ ,  $g(x_1,x_2)=(x_2,x_1)$ , then, if  $h(x_1,x_2)=f(g(x_1,x_2))$ :

- $\bigcirc$  a.  $\nabla h(x_1, x_2) = (2x_1, 2x_2)$ .
- $\bigcirc$  b.  $\nabla h(x_1, x_2) = (2x_2, 2x_1)$ .
- $\bigcirc$  c.  $\nabla h(x_1, x_2) = (1, 1)$ .

If  $f:\mathbb{R}^2 o\mathbb{R}$ ,  $f(x_1,x_2)=x_1e^{x_2}$ ,  $g:\mathbb{R} o\mathbb{R}^2$ ,  $g(t)=(t,\log t)$ , then, if h(t)=f(g(t)):

- $\circ$  a. h'(t) = t + 1.
- 0 b.  $h'(t) = t^2 + 1$ .
- $\bigcirc$  c. h'(t) = 2t.

If  $f:\mathbb{R}^2 o\mathbb{R}$ ,  $f(x_1,x_2)=x_1^2+x_1x_2$ ,  $g:\mathbb{R} o\mathbb{R}^2$ ,  $g(t)=(t^2,t)$ , then, if h(t)=f(g(t)):

$$oldsymbol{0}$$
 a.  $h'(t) = t(2t-1)^2 + t$ .

$$oldsymbol{0}$$
 b.  $h'(t) = 4t^2 + 2t + 1$ .

$$oldsymbol{\circ}$$
 c.  $h'(t) = t(2t+1)^2 - 2t^2$ .

If  $f:\mathbb{R}^2 o\mathbb{R}$ ,  $f(x_1,x_2)=x_1^2+x_2^2$ ,  $g:\mathbb{R}^2 o\mathbb{R}^2$ ,  $g(x_1,x_2)=(x_1e^{x_2},x_2)$ , then, if  $h(x_1,x_2)=f(g(x_1,x_2))$ :

- $\bigcirc$  a.  $\nabla h(x_1, x_2) = (2x_1e^{x_2}(e^{x_2} + x_1), 2e^{x_2}).$
- $\bigcirc$  b.  $\nabla h(x_1,x_2)=(2x_1e^{2x_2}(e^{x_1}+x_1),2x_2).$
- $\bigcirc$  c.  $\nabla h(x_1, x_2) = (2x_1e^{x_2}(e^{x_2} + x_1), 2x_2)$ .

If  $f:\mathbb{R}^2 o\mathbb{R}$ ,  $f(x_1,x_2)=x_1x_2$ ,  $g:\mathbb{R} o\mathbb{R}^2$ ,  $g(t)=(t,t^2)$ , then, if h(t)=f(g(t)):

- $\bigcirc$  a.  $h'(t)=3t^2$ .
- $\bigcirc$  b.  $h'(t) = 3t^3$ .
- $\bigcirc$  c.  $h'(t)=t^2$ .

lf

$$A = egin{bmatrix} 1 & 0 & 0 & 0 \ 0 & 1 & 0 & 0 \ 0 & 0 & 1 & 0 \ 0 & 0 & 0 & 1 \end{bmatrix}$$

then:

$$\bigcirc$$
 a.  $K_2(A)=1$ .

$$\bigcirc$$
 b.  $K_2(A)=4$ .

$$\bigcirc$$
 c.  $K_2(A)=rac{1}{2}$ .

$$A = egin{bmatrix} 2 & 0 & 0 & 0 \ 0 & 3 & 0 & 0 \ 0 & 0 & 2 & 0 \ 0 & 0 & 0 & 4 \end{bmatrix}$$

then:

- $\bigcirc$  a.  $K_2(A) = 4$ .
- $\bigcirc$  b.  $K_2(A) = 2$ .
- $\circ$  c.  $K_2(A) = \frac{4}{3}$ .

If vector  $v = (10^6, 0)^T$  is approximated by vector  $\tilde{v} = (999996, 1)^T$ , then in  $||\cdot||_2$  the relative error between v and  $\tilde{v}$  is:

- $\circ$  a.  $\sqrt{17} \cdot 10^{-6}$ .
- b. None of the above.
- $\bigcirc$  c.  $4 \cdot 10^{-6}$ .

$$A = egin{bmatrix} 2 & 0 & 0 & 0 \ 0 & 1 & 0 & 0 \ 0 & 0 & 2 & 0 \ 0 & 0 & 0 & 4 \end{bmatrix}$$

then:

$$\bigcirc$$
 a.  $K_2(A) = \frac{1}{2}$ .

Ob. 
$$K_2(A) = 4$$
.

$$\bigcirc$$
 c.  $K_2(A) = 2$ .

A random variable  $X:\Omega o \mathcal T$  is continuous when:

- $\bigcirc$  a.  $\mathcal{T}$  is countable.
- $\bigcirc$  b.  $\mathcal{T}=\mathbb{R}$ .
- $\odot$  c.  $\Omega$  is continuous.

A random variable  $X: \Omega \to \mathcal{T}$  is discrete when:

- $\bigcirc$  a.  $\mathcal{T} = \mathbb{R}$ .
- $\odot$  b.  $\Omega$  is countable.
- c. T is countable.

$$A = \left[egin{matrix} 4 & 2 \ 1 & 3 \end{matrix}
ight]$$

Then:

- $\bigcirc$  a.  $x=(1,2)^T$  is an eigenvector of A.
- $\bigcirc$  b.  $x=(2,1)^T$  is an eigenvector of A.
- $\bigcirc$  c.  $x=(0,0)^T$  is an eigenvector of A.

$$A = egin{bmatrix} 4 & 2 \ 1 & 3 \end{bmatrix}$$

## Then:

- igcup a.  $\lambda=5$  is the eigenvalue associated with the eigenvector  $x=(2,1)^T$  .
- igcup b.  $\lambda=2$  is the eigenvalue associated with the eigenvector  $x=(2,1)^T$  .
- $\odot$  c.  $\lambda=2$  is the eigenvalue associated with the eigenvector  $x=(1,2)^T$  .

$$A = egin{bmatrix} 4 & 0 & 0 \ 0 & 2 & 0 \ 0 & 0 & -1 \end{bmatrix}$$

# Then:

- $\bigcirc$  a.  $x=(1,1,0)^T$  is an eigenvector of A.
- $\bigcirc$  b.  $x=(0,1,0)^T$  is an eigenvector of A.
- $\bigcirc$  c.  $x=(0,-1,1)^T$  is an eigenvector of A.

$$A = egin{bmatrix} 4 & 0 \ 0 & 2 \end{bmatrix}$$

# Then:

- $\bigcirc$  a.  $x=(0,0)^T$  is an eigenvector of A.
- $\ \bigcirc$  b.  $x=(1,0)^T$  is an eigenvector of A.
- $\bigcirc$  c.  $x=(1,1)^T$  is an eigenvector of A.

lf

$$A = egin{bmatrix} 2 & 0 \ 0 & 1 \end{bmatrix}$$

Then:

- igcup a.  $\lambda=2$  is the eigenvalue associated with the eigenvector  $x=(1,0)^T$  .
- igcup b.  $\lambda=2$  is the eigenvalue associated with the eigenvector  $x=(0,1)^T$  .
- $\circ$  c.  $\lambda=1$  is the eigenvalue associated with the eigenvector  $x=(1,0)^T$  .

If  $A \in \mathbb{R}^{n \times n}$ ,  $x \in \mathbb{R}^n$  and

$$Ax = \lambda x$$

For  $\lambda \in \mathbb{R}$ , then:

- $\bigcirc$  a. For any  $c \in \mathbb{R}$ ,  $c \neq 0$ , cx is an eigenvector of A.
- $\bigcirc$  b. cx is an eigenvector of A if and only if c=1.
- c. None of the above.

$$A = egin{bmatrix} 2 & 0 & 0 \ 0 & 1 & 0 \ 0 & 0 & -1 \end{bmatrix}$$

## Then:

- $\bigcirc$  a.  $\lambda=2$  is the eigenvalue associated with the eigenvector  $x=(0,1,0)^T$  .
- $\odot$  b.  $\lambda=-1$  is the eigenvalue associated with the eigenvector  $x=(0,0,1)^T$ .
- $\odot$  c.  $\lambda=1$  is the eigenvalue associated with the eigenvector  $x=(1,0,0)^T$ .

In  $\mathcal{F}(10,2,-2,2)$ , if  $x=\pi$ , y=e, and z=fl(x)-fl(y), then:

- $\circ$  a.  $fl(z) = 0.43 \times 10^{1}$ .
- Ob.  $fl(z) = 0.44 \times 10^{1}$ .
- $\circ$  c.  $fl(z) = 0.40 \times 10^{1}$ .

In  $\mathcal{F}(10,2,-2,2)$ , if  $x=\pi$ , y=e, and z=fl(x)\*fl(y), then:

- $\circ$  a.  $fl(z) = 0.84 \times 10^{1}$ .
- $\circ$  b.  $fl(z) = 0.0837 imes 10^2$ .
- $\bigcirc$  c.  $fl(z) = 0.837 \times 10^1$ .

In  $\mathcal{F}(10,6,-3,3)$ , if x=192.403, y=0.635782, and z=fl(x)+fl(y), then:

- $\circ$  a.  $fl(z) = 0.193039 \times 10^3$ .
- $\circ$  b.  $fl(z) = 0.193038 \times 10^3$ .
- $\circ$  c.  $fl(z) = 0.193038782 \times 10^3$ .

In  $\mathcal{F}(10,2,-2,2)$ , if  $x=\pi$ , y=e, and z=fl(x)+fl(y), then:

- $\circ$  a.  $fl(z) = 0.585 \times 10^{1}$ .
- O b.  $fl(z) = 0.58 \times 10^{1}$ .
- $\circ$  c.  $fl(z) = 0.59 \times 10^{1}$ .

If  $f:\mathbb{R}^n o \mathbb{R}$ ,  $f \in \mathcal{C}^1(\mathbb{R}^n)$ , then  $x^*$  is a minimum point if and only if:

- igcirc a.  $abla f(x^*) = 0$  and  $abla^2 f(x^*)$  is positive semi-definite.
- $\bigcirc$  b.  $\nabla f(x^*) = 0$ .
- $\bigcirc$  c.  $abla f(x^*) = 0$  and  $abla^2 f(x^*)$  is positive definite.

## Gradient descent methods:

- $\square$  a. If  $\alpha$  is suitable chosen,  $f \in \mathcal{C}^1$ , for any  $x_0$ , always converges to a stationary point of f(x).
- b. If  $\alpha$  is suitable chosen,  $f \in C^1$ , for any  $x_0$ , always converges to a minimum of f(x).
- $\bigcirc$  c. Always converges to a minimum of f(x).

Gradient descent methods solves the optimization problem

$$\min_{x} f(x)$$

By:

- igcirc a. Generating a sequence  $\{x_k\}_k$  such that, given  $x_0$ , computes  $x_{k+1}=x_k-lpha
  abla f(x_k)$  for lpha>0 step-size.
- $\bigcirc$  b. Generating a sequence  $\{x_k\}_k$  such that, given  $x_0$ , computes  $x_{k+1}=x_k+lpha
  abla f(x_k)$  for lpha>0 step-size.
- $\odot$  c. Generating a sequence  $\{x_k\}_k$  such that, given  $x_0$ , computes  $x_{k+1}=x_k-lpha
  abla f(x_k)$  for lpha
  eq 0 step-size.

If  $f:\mathbb{R}^2 o\mathbb{R}$ ,  $f(x_1,x_2)=x_1^2+x_2^2$ , then if the initial guess for a gradient descent iteration is  $x^{(0)}=(1,1)^T$  and  $\alpha>0$ , then  $|f(x^{(1)})|<|f(x^{(0)})|$  if:

- 0 a.  $0 < \alpha < 1$ .
- $\bigcirc$  b.  $\alpha > 0$ .
- $\bigcirc$  c.  $\alpha > \frac{1}{2}$ .

If  $f:\mathbb{R}^2 \to \mathbb{R}$ ,  $f(x_1,x_2)=e^{x_1}+x_2^2$ , then if the initial guess for a gradient descent iteration is  $x^{(0)}=(0,0)^T$  and  $\alpha>0$ , then  $|f(x^{(1)})|<|f(x^{(0)})|$  if:

- $\bigcirc$  a.  $\alpha > \frac{1}{2}$ .
- $\bigcirc$  b.  $\alpha > 0$ .
- $\circ$  c.  $0 < \alpha < 1$ .

If  $f:\mathbb{R}^2\to\mathbb{R}$ ,  $f(x_1,x_2)=x_1e^{x_2}$ , then if the initial guess for a gradient descent iteration is  $x^{(0)}=(1,1)^T$  and  $\alpha=\frac{1}{2}$ , then:

$$\bigcirc$$
 a.  $x^{(1)}=(1-rac{e}{2},1-rac{e}{2})^T.$ 

$$\bigcirc$$
 b.  $x^{(1)} = (1 + \frac{e}{2}, 1 + \frac{e}{2})^T$ .

$$\bigcirc$$
 c.  $x^{(1)}=(rac{1}{2}-rac{e}{2},rac{1}{2}-rac{e}{2})^T.$ 

If  $f:\mathbb{R}^2 \to \mathbb{R}$ ,  $f(x_1,x_2)=x_1e^{x_2}$ , then if the initial guess for a gradient descent iteration is  $x^{(0)}=(0,0)^T$  and  $\alpha=1$ , then:

- $\bigcirc$  a.  $x^{(1)} = (1,0)^T$ .
- $\bigcirc$  b.  $x^{(1)} = (-1,0)^T$ .
- $\circ$  c.  $x^{(1)} = (0,0)^T$ .

If  $f:\mathbb{R}^2 o\mathbb{R}$ ,  $f(x_1,x_2)=x_1^2+x_2^2$ , then if the initial guess for a gradient descent iteration is  $x^{(0)}=(1,1)^T$  and lpha>0, then:

$$\bigcirc$$
 a.  $x^{(1)} = (1-2lpha, 1-2lpha)^T$ .

$$\circ$$
 b.  $x^{(1)} = (1 - \alpha, 1 - \alpha)^T$ .

$$\circ$$
 c.  $x^{(1)} = (1 + 2\alpha, 1 + 2\alpha)^T$ .

If  $f:\mathbb{R}^2 o\mathbb{R}$ ,  $f(x_1,x_2)=e^{x_1}+x_2^2$ , then if the initial guess for a gradient descent iteration is  $x^{(0)}=(0,0)^T$  and  $\alpha>0$ , then:

$$\bigcirc$$
 a.  $x^{(1)} = (-\alpha, 0)^T$ .

$$\bigcirc$$
 b.  $x^{(1)} = (0,0)^T$ .

$$\circ$$
 c.  $x^{(1)} = (-\alpha, 2)^T$ .

For Standard IEEE, double precision representation is:

- $\bigcirc$  a.  $\mathcal{F}(2,64,-1024,1023)$ .
- b. None of the above.
- $\bigcirc$  c.  $\mathcal{F}(2,53,-1024,1023)$ .

For Standard IEEE, single precision representation is:

- $\bigcirc$  a.  $\mathcal{F}(2,24,-128,127)$ .
- Ob. None of the above.
- $\bigcirc$  c.  $\mathcal{F}(2, 32, -128, 127)$ .

Given two independent random variables X and Y, then:

- $\bigcirc$  a. p(x) = p(y)
- $\bigcirc$  b. p(y) = p(y|x)
- $\bigcirc$  c. p(x|y) = p(y)

Given two random variables X and Y, Bayes Theorem implies that:

- $\bigcirc$  a. p(x) = p(y)p(y|x) / p(y|x).
- $\bigcirc$  b. p(x) = p(x|y)p(y|x) / p(y).
- $\bigcirc$  c. p(y) = p(y|x)p(x) / p(x|y).

If  $f:\mathbb{R}^n o\mathbb{R}$ ,  $f(x)=||Ax-b||_2^2$  for  $A\in\mathbb{R}^{m imes n},b\in\mathbb{R}^m$ , then:

$$\bigcirc$$
 a.  $abla f(x) = 2A^T(Ax-b)$ .

$$\bigcirc$$
 b.  $\nabla f(x) = A^T(Ax - b)$ .

$$\bigcirc$$
 c.  $\nabla f(x) = A(A^Tx - b)$ .

If  $f: \mathbb{R}^n \to \mathbb{R}$ ,  $f(x) = ||Ax - b||_2^2$  for  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ , then the solution of  $\nabla f(x) = 0$  is:

- $\bigcirc$  a.  $A^TAx = b$
- $\bigcirc$  b.  $A^TAx = A^Tx$
- $\bigcirc$  c. Ax = b.

# The machine precision $\epsilon$ can be defined as:

#### Select one:

- $\bigcirc$  a. The smallest number  $\epsilon$  such that  $fl(1+\epsilon)=1$ .
- $\bigcirc$  b. The smallest number  $\epsilon$  such that  $fl(1+\epsilon)>1$ .
- Oc. None of the above.

Given two random variables X and Y such that  $p(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x}$  and  $p(y|x)=ce^{-|y-ax|}$ , then the MAP reads:

$$\bigcirc \ \text{a.} \ x^* = \arg\min_x |y-ax| + \tfrac{1}{2}x^2.$$

$$\bigcirc$$
 b.  $x^* = \arg\min_x |y - ax|$ .

$$\bigcirc$$
 c.  $x^* = \arg\min_x \frac{1}{2}(y - ax)^2$ .

Given two random variables X and Y such that  $p(x)=ce^{-|x|}$  and  $p(y|x)=rac{1}{\sqrt{2\pi}}e^{-rac{1}{2}(y-ax)^2}$  , then the MAP reads:

- ${\mathbb Q}$  a.  $x^* = \arg\min_x \frac{1}{2} (y ax)^2 + |x|$ .
- $\bigcirc$  b.  $x^*=rg\min_xrac{1}{2}(y-ax)^2+rac{1}{2}x^2$ .
- $x^* = \arg\min_{x = \frac{1}{2}} (y ax)^2.$

Given two random variables X and Y such that  $p(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$  and  $p(y|x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y-ax)^2}$ , then the MAP reads:

$$\circ$$
 a.  $x^* = \arg\min_{x} \frac{1}{2} (y - ax)^2 + \frac{1}{2} x^2$ .

O b. 
$$x^* = \arg\min_{x \to 2} \frac{1}{2} (y - ax)^2 + x^2$$
.

$$x^* = \arg\min_{x} \frac{1}{2} (y - ax)^2$$
.

$$A = \left[ egin{array}{ccc} 2 & 1 & -2 \ -1 & 0 & 1 \ -1 & 2 & 1 \end{array} 
ight]$$

### Then:

- a. A is orthogonal.
- b. None of the above.
- c. A is symmetric and definite positive.

$$A = egin{bmatrix} 2 & 0 \ 0 & 1 \end{bmatrix}$$

Then:

- $\bigcirc$  a. A is symmetric and positive definite.
- $\bigcirc$  b. A is non-symmetric and not positive definite.
- $\bigcirc$  c. A is symmetric but not positive definite.

$$A = \left[ egin{array}{cccc} rac{2}{3} & rac{1}{3} & rac{2}{3} \ -rac{2}{3} & rac{2}{3} & rac{1}{3} \ rac{1}{3} & rac{2}{3} & -rac{2}{3} \end{array} 
ight]$$

## Then:

- $\bigcirc$  a. A is symmetric and definite positive.
- $\bigcirc$  b. A is symmetric but not definite positive.
- $\bigcirc$  c. A is orthogonal.

$$A = \left[ egin{array}{ccc} 2 & 2 & -1 \ 2 & 0 & 2 \ -1 & 2 & 3 \end{array} 
ight]$$

### Then:

- $\bigcirc$  a. A is symmetric but not definite positive.
- b. A is symmetric and definite positive.
- $\bigcirc$  c. A is orthogonal.

$$A = \left[ egin{array}{cc} -1 & 1 \ 0 & 3 \end{array} 
ight]$$

# Then:

- $\bigcirc$  a. A is symmetric but not positive definite.
- $\bigcirc$  b. A is symmetric and positive definite.
- $\odot$  c. A is non-symmetric and not positive definite.

$$A = \left[egin{matrix} 9 & 6 \ 6 & 5 \end{matrix}
ight]$$

# Then:

- a. A is symmetric but not positive definite.
- b. A is non-symmetric and not positive definite.
- c. A is symmetric and positive definite.

lf

$$A = \left[ egin{array}{cc} -1 & 0 \ 0 & 3 \end{array} 
ight]$$

# Then:

- $\bigcirc$  a. A is symmetric and positive definite.
- b. A is symmetric but not positive definite.
- $\bigcirc$  c. A is non-symmetric and not positive definite.

$$A = \left[ egin{array}{cccc} 2 & 1 & 2 \ -2 & 2 & 1 \ 1 & 2 & -2 \end{array} 
ight]$$

### Then:

- $\bigcirc$  a. A is not orthogonal.
- $\bigcirc$  b. A is orthogonal.
- $\bigcirc$  c. A is symmetric but not definite positive.

$$A = \left[ egin{array}{cccc} 1 & 0 & 3 \ 1 & 1 & 0 \ -1 & 1 & 2 \end{array} 
ight]$$

Then:

- $\bigcirc$  a. rank(A)=2.
- $\bigcirc$  b. rank(A) = 1.
- $\bigcirc$  c. rank(A)=3.

$$A = egin{bmatrix} 0 & 6 & 8 \ 2 & 4 & 0 \ 1 & 0 & 8 \end{bmatrix}$$

Then:

- $\bigcirc$  a. rank(A) = 3.
- $\bigcirc$  b. rank(A) = 2.
- $\bigcirc$  c. rank(A) = 1.

$$A = egin{bmatrix} 1 & 0 & 1 \ 0 & 1 & 0 \ 1 & 1 & 1 \end{bmatrix}$$

# Then:

- $\bigcirc$  a. rank(A) = 2.
- $\bigcirc$  b. rank(A) = 1.
- $\bigcirc$  c. rank(A)=3.

$$A = egin{bmatrix} 2 & 0 & 0 & 0 \ 0 & 3 & 0 & 0 \ 0 & 0 & 2 & 0 \ 0 & 0 & 0 & 0 \end{bmatrix}$$

then:

- $\bigcirc$  a. rank(A) = 3.
- $\bigcirc$  b. rank(A) = 4.
- $\bigcirc$  c. rank(A) = 2.

$$A = \left[ egin{array}{cccc} 1 & -1 & 1 \ 2 & -2 & 2 \ -1 & 1 & -1 \end{array} 
ight]$$

Then:

- $\bigcirc$  a. rank(A) = 3.
- $\bigcirc$  b. rank(A) = 1.
- $\bigcirc$  c. rank(A) = 2.

lf

$$A = egin{bmatrix} 0 & 2 & -1 \ 1 & 1 & 0 \ 1 & 3 & -1 \end{bmatrix}$$

# Then:

- $\bigcirc$  a. rank(A) = 2.
- $\bigcirc$  b. rank(A) = 3.
- $\bigcirc$  c. rank(A) = 1.

$$A = egin{bmatrix} 2 & 0 & 0 & 0 \ 0 & 3 & 0 & 0 \ 0 & 0 & 2 & 0 \ 0 & 0 & 0 & 4 \end{bmatrix}$$

then:

- $\bigcirc$  a. rank(A) = 4.
- $\bigcirc$  b. rank(A) = 2.
- $\bigcirc$  c. rank(A) = 3.

$$A = egin{bmatrix} 1 & 0 & 1 \ 0 & 1 & 0 \ 2 & 1 & 2 \end{bmatrix}$$

Then:

- $\bigcirc$  a. rank(A) = 1.
- $\bigcirc$  b. rank(A)=3.
- $\bigcirc$  c. rank(A) = 2.

Given two random variables X and Y such that  $p(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x}$  and  $p(y|x)=ce^{-|y-ax|}$ , then the MLE reads:

- $\bigcirc$  a.  $x^* = \operatorname{arg\,min}_x |y ax| + x^2$ .
- $\bigcirc$  b.  $x^* = rg \min_x rac{1}{2} (y-ax)^2$ .
- $\bigcirc$  c.  $x^* = \arg\min_x |y ax|$ .

Given two random variables X and Y such that  $p(x)=ce^{-|x|}$  and  $p(y|x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y-ax)^2}$ , then the MLE reads:

- $\bigcirc$  a.  $x^* = \arg\min_x \frac{1}{2} (y ax)^2 + |x|$ .
- Ob.  $x^* = \arg\min_x \frac{1}{2} (y ax)^2 + \frac{1}{2} x^2$ .
- $\odot$  c.  $x^* = \operatorname{arg\,min}_x \frac{1}{2} (y ax)^2$ .

Given two random variables X and Y such that  $p(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x}$  and  $p(y|x)=ce^{-|y-ax|}$  , then the MLE reads:

- $\circ$  a.  $x^* = \arg\min_x |y ax| + x^2$ .
- $\quad \ \ \, \text{b.} \ \ \, x^* = \arg\min_x |y-ax|.$
- $\circ$  c.  $x^* = \arg\min_{x} \frac{1}{2} (y ax)^2$ .

Given two random variables X and Y such that  $p(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$  and  $p(y|x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y-ax)^2}$ , then the MLE reads:

$$\bigcirc$$
 a.  $x^*=rg\min_x rac{1}{2}(y-ax)^2+rac{1}{2}x^2.$ 

$$\bigcirc$$
 b.  $x^* = rg \min_x rac{1}{2} (y-ax)^2 + x^2$ .

$$\bigcirc$$
 c.  $x^* = rg \min_x rac{1}{2} (y-ax)^2$ .

## then:

- $\bigcirc$  a. The 2-norm of A is  $||A||_2 = 1$ .
- $\bigcirc$  b. The 2-norm of A is  $||A||_2 = 0$ .
- $\odot$  c. The 2-norm of A is  $||A||_2 = 3$ .

$$A = egin{bmatrix} 2 & 0 & 0 & 0 \ 0 & 3 & 0 & 0 \ 0 & 0 & 2 & 0 \ 0 & 0 & 0 & 4 \end{bmatrix}$$

# then:

- $\bigcirc$  a. The 2-norm of A is  $||A||_2 = 2$ .
- $\bigcirc$  b. The 2-norm of A is  $||A||_2 = 4$ .
- $\odot$  c. The 2-norm of A is  $||A||_2 = 2$ .

$$A = egin{bmatrix} 1 & 0 & 0 & 0 \ 0 & 3 & 0 & 0 \ 0 & 0 & 2 & 0 \ 0 & 0 & 0 & 0 \end{bmatrix}$$

### then:

- $\bigcirc$  a. The 2-norm of A is  $||A||_2=3$ .
- $\bigcirc$  b. The 2-norm of A is  $\left|\left|A\right|\right|_2=0$ .
- $\odot$  c. The 2-norm of A is  $||A||_2 = 3$ .

If A is an n imes n matrix, then

### Scegli un'alternativa:

$$\bigcirc$$
 a.  $||A||_1=\sqrt{\sum_{i=1}^n\sum_{j=1}^na_{i,j}^2}.$ 

$$\bigcirc$$
 b.  $||A||_1 = 
ho(A^TA)$ .

oc. None of the above.

# If A is an n imes n matrix, then

- a. None of the above.
- $\bigcirc$  b.  $\left|\left|A\right|\right|_2=
  ho(A^TA).$
- $\bigcirc$  C.  $||A||_2 = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2}$ .

# If A is an n imes n matrix, then

$$\bigcirc$$
 a.  $||A||_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2}$ .

- Ob. None of the above.
- $||A||_F = \rho(A^T A).$

If  $A \in \mathbb{R}^{m imes n}$  ,  $||A||_p = 0$  , then:

- $\bigcirc$  a. A=0.
- $\bigcirc$  b. rank(A) = 0.
- $\odot$  c. A can be both equal or not equal to 0.

# A matrix $A \in \mathbb{R}^{n \times n}$ is orthogonal if:

$$\bigcirc \ \, {\rm a.} \ \ \, A^{-1}A = I = AA^{-1}.$$

$$\bigcirc \ \, \mathrm{b.} \quad A^TA = I = AA^T.$$

$$\bigcirc$$
 c.  $A=A^T$ .

If  $X:\Omega \to \mathcal T$  is a continuous random variable, then a function  $p:\mathcal T\to \mathbb R_+$  can be the PDF of X if:

- $\bigcirc$  a.  $\int_{\mathcal{T}} p(x) dx = 1$ .
- $\bigcirc$  b.  $\int_{\Omega}p(x)dx=1$ .
- $\bigcirc$  c.  $\int_{\mathcal{T}} p(x) dx < \infty$ .

For a random variable  $X:\Omega o \mathcal T$  with  $\mathbb E[X]=0$ , it holds:

- $\bigcirc$  a.  $Var(X) = \mathbb{E}[X]$ .
- ${\mathbb O}$  b.  $Var(X)={\mathbb E}[X^2].$
- $\bigcirc$  c. Var(X)=0.

If  $X:\Omega \to \mathcal{T}$  is a discrete random variable, then a function  $f_X:\mathcal{T} \to [0,1]$  can be the PDF of X if:

- $\bigcirc$  a.  $\sum_{i\in\Omega}f_X(i)=1$ .
- $\bigcirc$  b.  $\int_{\Omega}f_{X}(x)dx=1.$
- $\bigcirc$  c.  $\sum_{i\in\mathcal{T}}f_X(i)=1$ .

Given a discrete random variable  $X:\Omega o\mathcal T$ , with  $\mathcal T=\{1,2,\dots,6\}$ , and  $f_X=\{\frac16,\frac16,\dots,\frac16\}$ , then:

- $\bigcirc$  a.  $\mathbb{E}[X] = 21$ .
- $\odot$  b.  $\mathbb{E}[X] = 3.5$ .
- $\bigcirc$  c.  $\mathbb{E}[X] = \frac{1}{6}$ .

Given a continuous random variable  $X:\Omega \to \mathcal{T}$ , with  $\mathcal{T}=[0,1]$ , and  $p(x)=3x^2$  its PDF, then:

- ${\mathbb O}$  a.  ${\mathbb E}[X]=2$ .
- $\odot$  b.  $\mathbb{E}[X] = 3$ .
- $\bigcirc$  c.  $\mathbb{E}[X] = \frac{3}{4}$ .

Given a continuous random variable  $X:\Omega o \mathcal T$  , with  $\mathcal T=[0,1]$  , and p(x)=2x its PDF, then:

- $\bigcirc$  a.  $\mathbb{E}[X] = \frac{2}{3}$ .
- $\bigcirc$  b.  $\mathbb{E}[X] = 2$ .
- $\odot$  c.  $\mathbb{E}[X] = 1$ .

If  $X:\Omega o\mathcal{T}$  is a continuous random variable with PDG  $p:\mathcal{T} o\mathbb{R}_+$ , then:

$$\bigcirc$$
 a.  $\mathbb{E}[X] = \int_{\mathcal{T}} p(x) dx$ .

$$\odot$$
 b.  $\mathbb{E}[X] = \int_{\Omega} x p(x) dx$ .

$${\mathbb C}$$
 c.  ${\mathbb E}[X]=\int_{\mathcal T} x p(x) dx$ .

If  $X:\Omega o \mathcal T$  is a continuous random variable, its Probability Density Function (PDF)  $p_X(x)$  is defined to be:

$$\bigcirc$$
 a.  $P(X=x)=p_X(x)$ .

$$\bigcirc$$
 b.  $P(X=x)=\int x p_X(x) dx$ .

$$\bigcirc$$
 c.  $P(X \in A) = \int_A p_X(x) dx$ .

If  $X:\Omega\to\mathcal{T}$  is a continuous random variable, its Probability Density Function (PDF)  $p_X(x)$  is:

- $\bigcirc$  a. A function  $p_X: \mathcal{T} \to \mathbb{R}_+$ .
- $\bigcirc$  b. A function  $p_X: \mathcal{T} \to [0,1]$ .
- $\bigcirc$  c. A function  $p_X : \Omega \to [0, 1]$ .

If  $X:\Omega o \mathcal T$  is a discrete random variable, its Probability Mass Function (PMF)  $f_x$  is:

#### Select one:

- igcirc a. A function  $f_X:\mathcal{T} o [0,1].$
- ${}$  b. A function  $f_X:\Omega o [0,1].$
- ${\mathbb C}$  c. A function  $f_X:{\mathcal T} o{\mathbb R}$  .

Given two discrete random variable  $X_1:\Omega o\mathcal{T}$ ,  $X_2:\Omega o\mathcal{T}$  with  $\mathcal{T}=\{1,2,3\}$ , and  $f_{X_1}=\{\frac{1}{3},\frac{1}{3},\frac{1}{3}\}$ ,  $f_{X_2}=\{\frac{1}{2},\frac{1}{6},\frac{1}{3}\}$  their PMF, then:

- $\odot$  a.  $\mathbb{E}[X_1] < \mathbb{E}[X_2]$ .
- $\bigcirc$  b.  $\mathbb{E}[X_1] = \mathbb{E}[X_2]$ .
- $\odot$  c.  $\mathbb{E}[X_1] > \mathbb{E}[X_2]$ .

If  $X:\Omega \to \mathcal{T}$  is a discrete random variable, its Probability Mass Function (PMF)  $f_x$  is:

- $\bigcirc$  a.  $f_X(x) = \int P(x) dx$ .
- $\bigcirc$  b.  $f_X(x)=P(X\in x)$ .
- $\bigcirc$  c.  $f_X(x) = P(X = x)$ .

If  $X:\Omega o\mathcal{T}$  is a discrete random variable with PMG  $f_X:\mathcal{T} o[0,1]$ , then:

#### Scegli un'alternativa:

$$\odot$$
 a.  $\mathbb{E}[X] = \sum_{i \in \mathcal{T}} i$ .

igtherightarrow b.  $\mathbb{E}[X] = \sum_{i \in \mathcal{T}} i f_X(i)$ .

Vai a...

Given a discrete random variable  $X:\Omega\to\mathcal{T}$ , with  $\mathcal{T}=\{1,2,3\}$ , and  $f_X=\{\frac{1}{2},\frac{1}{6},\frac{1}{3}\}$  its PMF, then:

- $\bigcirc$  a.  $\mathbb{E}[X] = 6$ .
- $\odot$  b.  $\mathbb{E}[X]=2$ .
- $\bigcirc$  c.  $\mathbb{E}[X] = \frac{11}{6}$ .

Given a discrete random variable  $X:\Omega o\mathcal T$ , with  $\mathcal T=\{1,2,3\}$ , and  $f_X=\{\frac16,\frac13,\frac12\}$  its PMF, then:

- $\bigcirc$  a.  $\mathbb{E}[X] = 6$ .
- $\odot$  b.  $\mathbb{E}[X]=2$ .
- $\bigcirc$  c.  $\mathbb{E}[X] = \frac{13}{6}$ .

Given two discrete random variable  $X_1:\Omega\to\mathcal{T}$ ,  $X_2:\Omega\to\mathcal{T}$  with  $\mathcal{T}=\{1,2,3\}$ , and  $f_{X_1}=\{\frac{1}{3},\frac{1}{3},\frac{1}{3}\}$ ,  $f_{X_2}=\{\frac{1}{2},\frac{1}{6},\frac{1}{3}\}$  their PMF, then:

- $\odot$  a.  $\mathbb{E}[X_1] < \mathbb{E}[X_2]$ .
- ${\mathbb O}$  b.  ${\mathbb E}[X_1] = {\mathbb E}[X_2].$
- $\odot$  c.  $\mathbb{E}[X_1] > \mathbb{E}[X_2]$ .

Given two discrete random variable  $X_1:\Omega\to\mathcal{T}$ ,  $X_2:\Omega\to\mathcal{T}$  with  $\mathcal{T}=\{1,2,3\}$ , and  $f_{X_1}=\{\frac12,\frac16,\frac13\}$ ,  $f_{X_2}=\{\frac16,\frac13,\frac12\}$  their PMF, then:

- ${\mathbb C}$  a.  ${\mathbb E}[X_1] > {\mathbb E}[X_2]$ .
- $\odot$  b.  $\mathbb{E}[X_1] = \mathbb{E}[X_2]$ .
- $\odot$  c.  $\mathbb{E}[X_1] < \mathbb{E}[X_2]$ .

# The precision is:

- a. None of the above.
- $\bigcirc$  b. The number of digits with which a number is expressed.
- o. The number of correct significant digits in approximating some quantity.

# A random variable X is:

- $\bigcirc$  a. A function  $X:\Omega o \mathcal{T}$ .
- b. A variable that returns random elements with known probability.
- c. A set that contains the possible outcomes of the experiment.

If  $\Omega$  is the sample space,  ${\mathcal A}$  is the event space and  ${\mathcal T}$  is a subset of  ${\mathbb R}$ , a random variable X is:

- $\bigcirc$  a. A function  $X:\Omega \to \mathcal{A}$ .
- $\bigcirc$  b. A function  $X: \mathcal{A} \rightarrow \mathcal{T}$ .
- $\bigcirc$  c. A function  $X:\Omega 
  ightarrow \mathcal{T}$ .

In normalized scientific notation and base  $\beta=10$ , if x=2.71, then:

- $\odot$  a. The mantissa of x is 0.271 and the exponential part is  $10^1$ .
- $\odot$  b. The mantissa of x is 2.71 and the exponential part is  $10^0$ .
- oc. None of the above.

If  $A=U\Sigma V^T$  is the SVD decomposition of  $A\in\mathbb{R}^{m imes n}$  , then a dyade  $A_i=u_iv_i^T$  of A is:

- a. None of the above.
- $\bigcirc$  b. A vector of length mn that express some properties of A.
- $\bigcirc$  c. A rank-1 matrix of dimension  $m \times n$ .

If  $A = U \Sigma V^T$  is the SVD decomposition of an m imes n matrix A, then:

#### Scegli un'alternativa:

$$\bigcirc$$
 a.  $A^TA = V^T\Sigma^2V$ .

$$\bigcirc$$
 b.  $A^TA = U\Sigma^2U^T$ .

 $\bigcirc$  c.  $A^TA=V\Sigma^2V^T$ .

If  $A = U\Sigma V^T$  is the SVD decomposition of an  $m \times n$  matrix A, then:

- $\bigcirc$  a. The rows of  $V^T$  are eigenvectors of  $AA^T$ .
- $\bigcirc$  b. The columns of U are eigenvectors of  $AA^T$ .
- o. None of the above

If  $A = U \Sigma V^T$  is the SVD decomposition of an m imes n matrix A, then:

- $\odot$  a. The rows of  $V^T$  are eigenvectors of  $A^TA$ .
- b. None of the above
- $\odot$  c. The columns of U are eigenvectors of  $A^TA$ .

Given a matrix  $A \in \mathbb{R}^{m \times n}$  , m > n , with r = rank(A) , then:

- a. None of the above.
- $\odot$  b. It is always possible to write A as  $U\Sigma V^T$ , where  $\Sigma\in\mathbb{R}^{m\times n}$  is diagonal,  $U\in\mathbb{R}^{m\times m}$ ,  $V\in\mathbb{R}^{n\times n}$  are orthogonals.
- $\odot$  c. It is possible to write A as  $U\Sigma V^T$ , where  $\Sigma\in\mathbb{R}^{m\times n}$  is diagonal,  $U\in\mathbb{R}^{m\times m}$ ,  $V\in\mathbb{R}^{n\times n}$  are orthogonals if and only if rank(A)=n.

If  $A=U\Sigma V^T$  is the SVD decomposition of an m imes n matrix A, then:

- $\bigcirc$  a.  $AA^T=U^T\Sigma^2U$ .
- $\bigcirc$  b.  $AA^T=U\Sigma^2U^T$ .
- $\bigcirc \ \, \mathsf{c}, \quad AA^T = U\Sigma U^T.$

If  $A=U\Sigma V^T$  is the SVD decomposition of an m imes n matrix A, then:

- igcirc a. The elements of the diagonal matrix  $\Sigma$  are the singular values of A, in decreasing order.
- $\odot$  b. The elements of the diagonal matrix  $\Sigma$  are the eigenvalues of A, in decreasing order.
- o. None of the above.

If  $A=U\Sigma V^T$  is the SVD decomposition of an m imes n matrix A, then:

- $\bigcirc$  a. The singular values  $\sigma_i$  of A are  $\sigma_i = \sqrt{\lambda_i(A^TA)}$  where  $\lambda_i(A^TA)$  are the eigenvalues of  $A^TA$ .
- b. None of the above
- $\odot$  c. The singular values  $\sigma_i$  of A are  $\sigma_i = \lambda_i(A^TA)$  where  $\lambda_i(A^TA)$  are the eigenvalues of  $A^TA$ .

If  $A=U\Sigma V^T$  is the SVD decomposition of  $A\in\mathbb{R}^{m imes n}$  , then the rank k approximation of A is:

- igcirc a.  $\hat{A}(k) = \sum_{i=1}^n \sigma_i A_i$  , where  $A_i = u_i v_i^T$  is a dyade.
- igcirc b.  $\hat{A}(k) = \sum_{i=1}^k A_i$  , where  $A_i = u_i \, v_i^T$  is a dyade.
- igcup c.  $\hat{A}(k) = \sum_{i=1}^k \sigma_i A_i$  , where  $A_i = u_i v_i^T$  is a dyade.

If  $A = U \Sigma V^T$  is the SVD decomposition of an m imes n matrix A, then:

- $\odot$  a. The elements on the diagonal of  $\Sigma$  are strictly positive.
- b. None of the above.
- $\odot$  c. The elements of the diagonal matrix  $\Sigma$  are non-negative.

If  $A=U\Sigma V^T$  is the SVD decomposition of  $A\in\mathbb{R}^{m imes n}$ , then its rank k approximation of  $\hat{A}(k)$  satisfies:

#### Select one:

- $\bigcirc$  a.  $\hat{A}(k) = rg \min_{rk(B)=k} ||A-B||_2$ .
- $\bigcirc$  b.  $\hat{A}(k) = rg \min_{rk(B)=k} ||A-B||_F$ .
- $\bigcirc$  c.  $\hat{A}(k)=\sigma_{k+1}$ .

If  $A \in \mathbb{R}^{m \times n}$ , then:

- $\bigcirc$  a.  $A^TA$  is symmetric but not necessarly positive definite.
- $\bigcirc$  b.  $A^TA$  is always symmetric and positive definite.
- $\odot$  c. It depends on A.

For a random variable  $X:\Omega\to\mathcal{T}$ , its variance is defined as:

- $\bigcirc$  a.  $Var(X) = \mathbb{E}[X \mathbb{E}[X]].$
- $\bigcirc$  b.  $Var(X) = \mathbb{E}[(X \mathbb{E}[X])^2].$
- $\bigcirc$  c.  $Var(X) = \mathbb{E}[X^2 \mathbb{E}[X]^2]$ .

For a random variable  $X:\Omega \to \mathcal{T}$ , it holds:

- igcirc a.  $Var(X)=\mathbb{E}[X^2]-\mathbb{E}[X]^2$ .
- $\odot$  b.  $Var(X) = \mathbb{E}[X^2] + \mathbb{E}[X]^2$ .
- ${\mathbb C}$  c.  $Var(X)={\mathbb E}[X]^2-{\mathbb E}[X^2]$ .

For a random variable  $X:\Omega o\mathcal{T}$  with  $\mathbb{E}[X]=0$ , it holds:

- $\bigcirc$  a.  $Var(X)=\mathbb{E}[X^2].$
- b. Var(X) = 0.
- $\bigcirc$  c.  $Var(X) = \mathbb{E}[X]$ .

If vector  $v=(10^6,0)^T$  is approximated by vector  $\tilde{v}=(999996,1)^T$ , then in  $||\cdot||_1$  the relative error between v and  $\tilde{v}$  is:

- $\bigcirc$  a.  $4 \cdot 10^{-6}$ .
- b. None of the above.
- $\circ$  c.  $5 \cdot 10^{-6}$ .

If vector  $v=(10^6,0)^T$  is approximated by vector  $\tilde{v}=(999996,1)^T$ , then in  $||\cdot||_2$  the relative error between v and  $\tilde{v}$  is:

- $\bigcirc$  a.  $\sqrt{17} \cdot 10^{-6}$ .
- O b.  $4 \cdot 10^{-6}$ .
- c. None of the above.

If vector  $v=(1,0)^T$  is approximated by vector  $\tilde{v}=(1.1,0.1)^T$ , then in  $||\cdot||_1$  the relative error between v and  $\tilde{v}$  is:

- a. 0.1.
- O b. 0.2.
- o. None of the above.

If vector  $v=(10^6,1)^T$  is approximated by vector  $\tilde{v}=(999996,1)^T$ , then in  $||\cdot||_\infty$  the relative error between v and  $\tilde{v}$  is:

- O a. 4.
- $\circ$  b.  $4 \cdot 10^{-6}$ .
- c. None of the above.

Given  $x=(1,-1)^T$  and  $y=(2,1)^T$  , then:

- a. None of the above.
- $\bigcirc$  b.  $x \perp y$ .
- $\bigcirc$  c.  $x \parallel y$ .

Given  $x = (1, -1)^T$  and  $y = (1, 1)^T$ , then:

- $\bigcirc$  a.  $x \perp y$ .
- b. None of the above.
- $\bigcirc$  c.  $x \parallel y$ .